ESDCD news

Winter 2005

Code 930 • http://esdcd.gsfc.nasa.gov

Earth and Space Data Computing Division

Earth Sciences Directorate, Goddard Space Flight Center

In This Issue

ESDCD Updates

Applications Take the Floor at SC2004 Conference 1

GSFC Among First 10 Users of the National LambdaRail 3

Interstellar Sugar Discovery Makes *Discover* Magazine's 100 Top Science Stories of 2004 6

NCCS Highlights

NCCS Welcomes a New SGI Altix High-Performance Computer 6

Computational Technologies Project

Second ESMF IAWG Meeting Held at GSFC 7

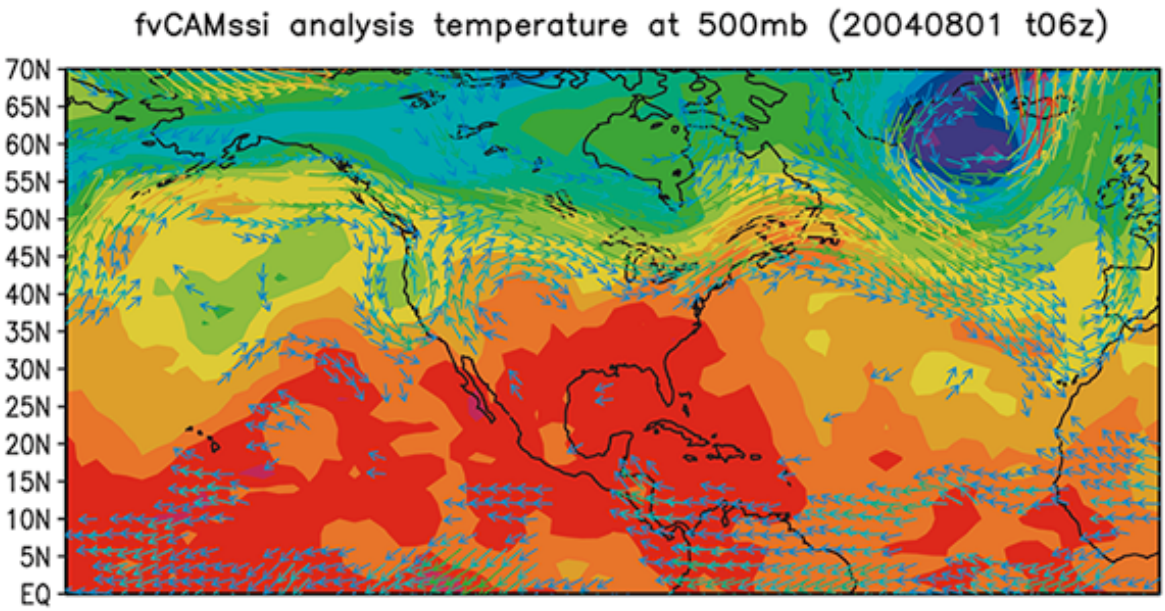
ESDCD Updates

Applications Take the Floor at SC2004 Conference

"Preparing for the SC conference is one of the driving forces to push work to the next level," said Shujia Zhou of the Computational Technologies (CT) Project and Northrop Grumman IT. "It is not like a traditional academic conference. Here, you try to demonstrate something that is working and that outsiders can understand." Zhou presented one of seven CT- and ESDCD-affiliated demonstrations of supercomputing applications in NASA's research exhibit at SC2004, the world's leading high-performance computing (HPC), networking, and storage conference. SC2004 drew a record international audience of nearly 8,000 people to Pittsburgh, PA, November 6–12.

Zhou demonstrated recent developments with the Earth System Modeling Framework (ESMF) (see "Second ESMF IAWG Meeting Held at GSFC" in this issue). In October, researchers completed a field test using ESMF to couple a Goddard Space Flight Center (GSFC)-National Center for Atmospheric Research (NCAR) atmosphere model with the National Centers for Environmental Prediction's (NCEP) operational data assimilation technology for the first time. "When we prepared for the ESMF demo, we had to do something beyond the mechanical coupling of two components," Zhou explained. "It is along the lines of something with scientific merit." The combination of the GSFC-NCAR finite-volume Community Atmosphere Model (fvCAM) with the NCEP Spectral Statistical Interpolation (SSI) Analysis System produces global temperature and wind output reasonably consistent with that from the NCEP operational system used in weather forecasting. Zhou also showed a grid-enabled version of the ESMF-Common Component Architecture prototype that had been run on two partitions of the ESDCD's Thunderhead cluster across a newly implemented 10 gigabit-per-second network. This implementation was part of the ESDCD National LambdaRail effort (see "Goddard Space Flight Center Among First 10 Users of the National LambdaRail" in this issue).

The NASA research exhibit (center) was one of 266 industry and research exhibits at the SC2004 conference, which attracted record attendance of nearly 8,000 people (Photo credit: Jarrett Cohen, Global Science and Technology, Inc.).



A field test of the Earth System Modeling Framework coupled a GSFC-National Center for Atmospheric Research (NCAR) atmosphere model (fvCAM) with National Centers for Environmental Prediction (NCEP) operational data assimilation technology (SSI) for the first time. This visualization shows temperature and wind output from the fvCAM-SSI coupled system; temperature increases with reddening. Researchers on this effort were Carlos Cruz and Shujia Zhou, Northrop Grumman IT; Weiyu Yang, NCEP; Erik Kluzek, NCAR; and Arlindo da Silva, GSFC Global Modeling and Assimilation Office (Image credit: Carlos Cruz).

A GSFC atmosphere model that successfully forecasted aspects of 2004 Atlantic Ocean hurricanes was the subject of a demonstration as well as a Technical Program lecture at SC2004. Results from finite-volume General Circulation Model (fvGCM) hurricane simulations were shown by Bill Putman of the ESDCD, Bo-Wen Shen of Science Applications International Corporation, and Jiun-Dar Chern of the University of Maryland, Baltimore County, with on-site visualization support from Randall Jones of Global Science and Technology, Inc. Putman presented the fvGCM in a special Birds-of-a-Feather session on "Late Breaking Results," which featured highlights from the two fastest supercomputers in the world—the Department of Energy's Blue Gene/L and NASA's Columbia. When introducing Putman, session chair Dave Cooper of the Lawrence Livermore National Laboratory said the fvGCM research produced "some of the most fascinating simulation results I have ever seen."

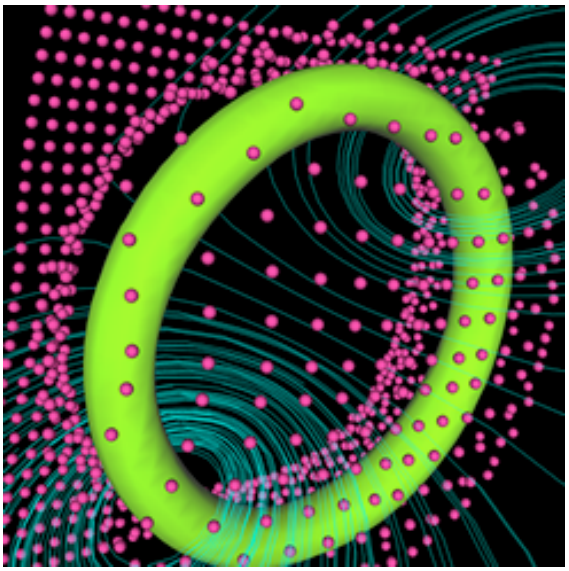
CT investigators demonstrated software frameworks for simulating space science applications of importance to NASA. IBEAM focuses on gamma-ray bursts, colossal explosions that can outshine one billion suns. "This is the first time that radiation flow and fluid flow have been coupled," said Doug Swesty of the State University of New York at Stony Brook. Such realism enables comparing simulated light curves with those observed by NASA space telescopes. IBEAM also incorporates several computer science innovations that run on adaptive meshes. "The way we solve the equations is not the way people normally do it," Swesty stressed. "Traditional nuclear engineering approaches do not scale to thousands of processors. Our approach was developed just in the last year or two and is now starting to be seen as the way to solve the problem."

The Chombo framework is a set of building blocks that space scientists can combine in specific ways, using adaptive mesh refine-

ment (AMR) to follow multi-scale changes in space and time. “Chombo minimizes the pain in making your application into an AMR application,” said Dan Martin of the Lawrence Berkeley National Laboratory. Collaborators at NASA’s Glenn Research Center plan to use Chombo for microgravity research. As Martin explained, the scientists “want a code that can model suspended particles in an incompressible fluid: the particles exert drag on the flow and vice versa.” In a representative problem shown at SC2004, small particles are suspended in water, such as might occur in microgravity mixing processes. The simulated microgravity environment generates a vortex ring of rotating fluid, a phenomenon that can be matched against real-world measurements.

Clemson University’s Parallel Architecture Research Laboratory (PARL), a long-time CT grant recipient, showed their latest software for making advanced HPC systems usable by non-specialist science and engineering professionals. They announced PVFS2 version 1.0, a parallel file system for Linux clusters that provides concurrent access to data distributed across a collection of servers. A multi-cluster computational grid simulator called BeoSim is being used to explore the effectiveness of multi-site scheduling algorithms under a variety of workload characterizations and grid configurations. Director Walt Ligon and graduate students Will Jones, Louis Pang, Brad Settlemyer, and Mike Speth represented PARL at SC2004.

Jet Propulsion Laboratory scientists Chris Catherasoo and Heidi Lorenz-Wirzba demonstrated the work of CT investigations based there and at Caltech. They explained how the QuakeSim investigation’s GeoFEST software was parallelized with the PYRAMID AMR library and ported to Columbia and other HPC systems. The researchers also showed simulations of earthquake processes, particularly those involving California’s Landers and Los Angeles basin faults. Space science demonstrations involved the



In a simulation from the Chombo software framework, a vortex ring interacts with suspended drag particles. Coherent vortices such as that shown here are a principal means for mixing fluids in microgravity environments, including the orbiting Space Shuttle and International Space Station (Image credit: Dan Martin, Lawrence Berkeley National Laboratory).

Montage software, which produces seamless astronomical image mosaics for the National Virtual Observatory from input data sets tens of terabytes in size. Researchers described Montage’s ability to run in computing environments ranging from a PC (where mosaic production takes hours) to 128 processors of the National Science Foundation’s TeraGrid (where production takes seconds).

http://www.sc-conference.org/sc2004

GSFC Among First 10 Users of the National LambdaRail

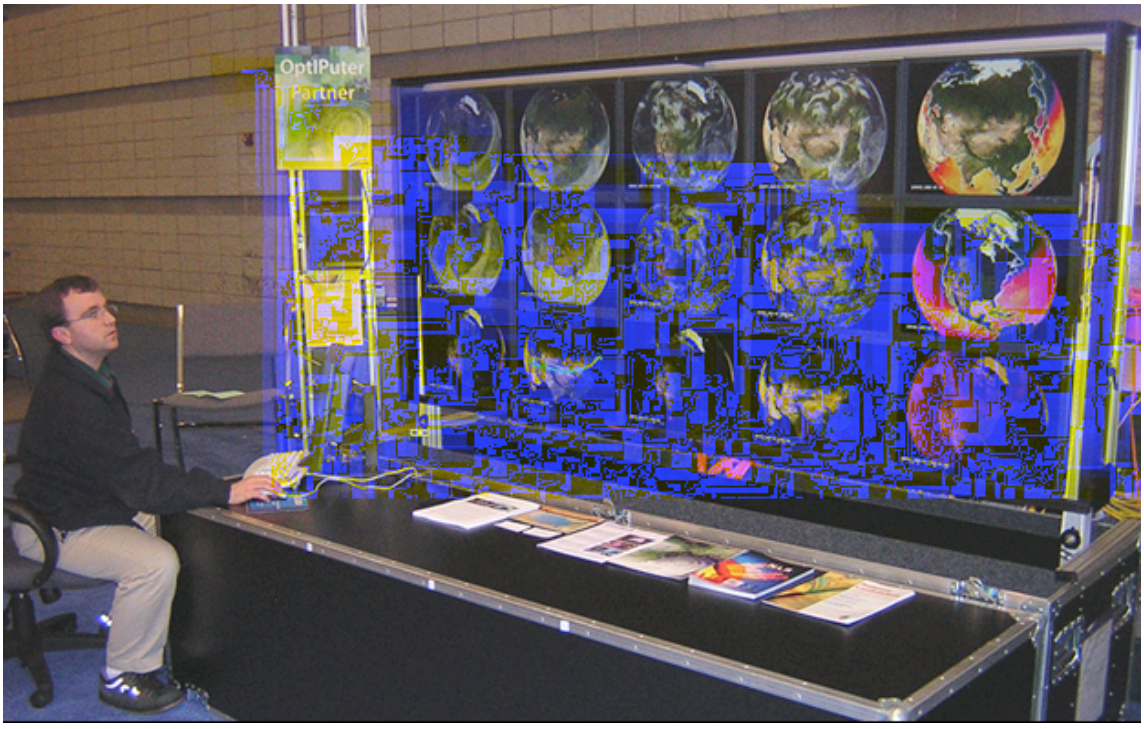
Goddard Space Flight Center (GSFC) is one of the first 10 sites to send scientific data over the National LambdaRail (NLR), a growing multi-wavelength optical network that is linking research sites across the United States, presently at 10 gigabits per second (Gbps) per wavelength. With 40 times more bandwidth and number of wavelengths than the current national re-

The National LambdaRail (NLR) is a growing multi-wavelength optical network that is linking research sites across the United States at 10 gigabits per second. This diagram shows Phase 1 of the NLR (Image credit: Debbie Montano, NLR).

search network, the NLR will allow high-performance computers and data facilities hundreds of miles apart to communicate as if they were in nearby buildings.

A lambda is a small range of light wavelengths; multiple lambdas can co-exist on a single optical fiber. Lambda networks exploit "dark" (laid down but previously unused) fiber lines, using optical switches to route packets of light rather than convert them to electronic signals. The NLR's nationwide backbone currently has four lambdas that stretch from Jacksonville, FL; northward to Washington, DC; then westward to Seattle, WA; and finally southward to San Diego, CA.

A GSFC Internal Research and Development (IRAD) project is funding the initial connection of GSFC high-performance computing resources to the NLR. "We want to position GSFC so it can be part of the broader science world, so GSFC can be competitive," said the ESDCD's Pat Gary, who is IRAD co-principal investigator (PI) with Jeff Smith of the Information Systems Division (ISD). Gary leads the High-End Computer Network team, which is implementing the NLR connection and 10 Gbps Ethernet across GSFC. Team members include Bill Fink of the ESDCD; Kevin Kranacs

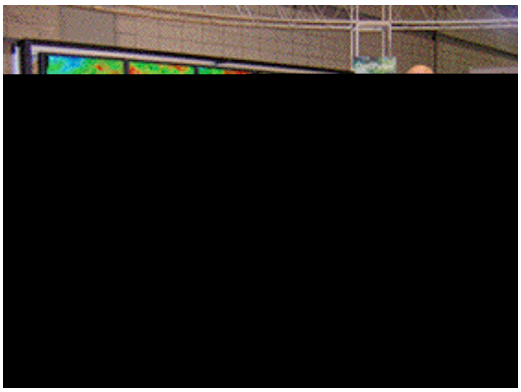


GSFC's connection to the NLR sent Earth science data sets in real time to a 15-screen tiled display at the SC2004 conference in Pittsburgh, PA. Eric Sokolowsky of the Scientific Visualization Studio (SVS) and Global Science and Technology (GST), Inc., interactively viewed model and observation data from the SVS's Image Server project (Photo credit: Randall Jones, GST, Inc.).

of the ISD; Paul Lang and Aruna Muppalla of ADNET Systems, Inc.; and Jeff Martz and Mike Stefanelli of Computer Sciences Corporation.

GSFC's NLR link debuted at the SC2004 high-performance computing, networking, and storage conference in Pittsburgh, PA, this past November (see "Applications Take the Floor at SC2004 Conference" in this issue). In cooperation with the National Science Foundation (NSF)-funded OptIPuter Project, interactive demonstrations flowed multiple Earth science data sets in real time to a 15-screen tiled display in the NLR research exhibit. "The involvement of NASA Goddard demonstrated the capabilities of NLR and showed just how researchers in 'big science' will need this kind of capacity to make new discoveries about aspects of our world and to help transfer this knowledge to practical uses by others in carrying out important tasks that improve our lives," said Tom West, president and CEO of the NLR.

At SC2004, Scientific Visualization Studio (SVS) programmers Randall Jones and Eric Sokolowsky, both of Global Science and Technology, Inc., previewed applications that NLR access will enable at GSFC. A demonstration of the SVS's Image Server project pulled data residing on servers in McLean, VA; Chicago, IL; and San Diego, CA to Pittsburgh. The data included output from high-resolution NASA atmosphere and ocean models and observations from satellites such as Aura, Polar, and the Tropical Rainfall Measuring Mission. Using the SVS-developed Digital Earth PC software, Sokolowsky said he could "interactively explore the scientific imagery using a 3-D model of the globe," as well as easily switch sets of images. "The audience was especially interested in watching 15 different views being controlled simultaneously," he said. Jones presented 1-km global data sets from the Computational Technologies Project's Land Information System (LIS) investigation. Pulling data from servers in Chicago and San Diego, and across the Atlantic Ocean in



At SC2004, Randall Jones of the SVS and GST, Inc., showed output from the Land Information System, which models the global land surface at 1-km resolution. The synchronized views compare U.S. surface evaporation with Mexico surface temperature (Photo credit: Eric Sokolowsky, GST, Inc.).

Amsterdam, Jones synchronized views of different countries and regions using various combinations of displays. The 1-km global LIS data sets range up to 100 gigabytes and span 40,000 x 20,000 cells, attesting to the need for ultra-high-speed networks to move them as well as tiled displays that can show far more of the data than individual workstation displays.

Back at GSFC, tiled displays in the SVS and the LIS laboratory will soon be linked with 10 Gbps Ethernet, equaling data transmission rates from the NLR. According to SVS Manager Horace Mitchell, GSFC Earth scientists listed tiled displays as significant collaborative technology because of the ability to simultaneously display multiple registered parameters. Resources already connected with 10 Gbps Ethernet include the Thunderhead cluster (in two partitions for grid emulation), the Goddard Geophysical and Astronomical Observatory, and several test workstations.

GSFC's near-term NLR plans call for incorporating high-performance computing, visualization, and storage technologies into the OptIPuter, linking GSFC facilities with those at the Scripps Institution of Oceanography; the University of California, San Diego (UCSD); and the University of Illinois at Chi-

cago (UIC). GSFC and Scripps are particularly keen on strengthening and expanding collaborations in aerosols, the global water and energy cycles, land cover, topography, and other areas. Integrating resources at these far-flung locations is also an important step in the build-up of the LambdaGrid envisioned by OptIPuter PI Larry Smarr of UCSD and Co-PI Tom DeFanti and Project Manager Maxine Brown of UIC, with whom GSFC is closely partnering. "The OptIPuter's lambda is known as the CAVEwave. It is essentially a fifth coast-to-coast lambda on the NLR," Gary said. "GSFC is an extension of the CAVEwave and part of the first coast-to-coast private lambda."

GSFC is NASA's first Center using the NLR and will be joined by other NASA Centers later this year. As of January 2005, GSFC computers connected to the NLR are located in the NLR suite at the Level3 Communications' optical fiber "carrier hotel" facility in McLean, VA. In early March, two 10 Gbps lambdas will be enabled across DRAGON, a Washington, DC-area multi-wavelength research network that is funded by NSF. These lambdas will link McLean with resources at GSFC's main site in Greenbelt, MD.

http://esdcd.gsfc.nasa.gov/IRAD_Lambda.html
http://www.nlr.net

Interstellar Sugar Discovery Makes *Discover* Magazine's 100 Top Science Stories of 2004

Congratulations to the ESDCD's Jan M. Hollis and his research team for making *Discover* magazine's 100 Top Science Stories of 2004 (January 2005 issue). *Discover* ranked their detection of two aldehyde molecules, propenal and propanal, in the cold interstellar cloud Sagittarius B2 as #26 for its implications for the molecular origins of life.

http://www.discover.com/issues/jan-05/features/top-100stories/
http://esdcd-news.gsfc.nasa.gov/2004.Fall/07_sugar_space.html
http://www.nasa.gov/vision/universe/starsgalaxies/interstellar_sugar.html

NCCS Highlights

NCCS Welcomes a New SGI Altix High-Performance Computer

The NCCS recently acquired an SGI Altix 3000 high-performance computer to enable users to port and tune their code and perform limited scaling tests for applications in an environment like that of NASA Ames Research Center's Project Columbia.

The new Altix, named in honor of Nancy L. Palm, former Head of the Science Computing Branch of the ESDCD, has 128 processors, 768 gigaFLOPS peak computing capacity, and 286 gigabytes of total memory. Its 17 terabytes of raw serial advanced technology attachment (SATA) disk sit behind a TP9300S controller, and extra I/O bricks are available to break the system into at least two images using an InfiniBand switch and host card adapters.

Initial tests on the new system have been performed with the NASA finite-volume General Circulation Model (fvGCM). When using 120 processors, the new system has been producing throughput on the order of

The NASA Center for Computational Science's new SGI Altix 3000 high-performance computer will complement those of NASA Ames Research Center's Project Columbia (Photo credit: NASA Ames Research Center/Tom Trower).

15 model months per wall-clock day for a climate application of the 1/2-degree, 32-level atmospheric model. Standard 10-day numerical weather prediction experiments with the fvGCM at 1/2-degree resolution, which typically take 1 hour to complete with 128 processors on the HP/Compaq SC45, can be performed in half the time on the new system.

After the machine is fully tested and incorporated into NCCS User Support Group processes, the NCCS will allow users to request dedicated time on the system. The Altix and other NCCS SGI environments will subsequently be integrated into one CXFS fabric, and the possibility exists for bringing a much larger Altix system to the NCCS to complement that of Columbia.

Training on the new Altix, hosted by the NCCS, SGI, and Intel, will feature an overview of the Altix system and of Columbia; an introduction to the architecture of the Altix system, Itanium processor, and Intel compilers and tools; and lessons in tuning and scaling.

Computational Technologies Project

Second ESMF IAWG Meeting Held at GSFC

The Earth System Modeling Framework (ESMF) Interagency Working Group (IAWG) held its second biannual meeting at GSFC on November 30, 2004. ESMF is a national-scale collaboration to build a software infrastructure that allows weather and climate model components from different researchers to operate together on parallel supercomputers. ESMF has garnered support at high levels because it fills a variety of needs such as enabling long-term cost savings and facilitating cross-agency collaboration.

The ESMF IAWG was called into being by NASA Headquarters to coordinate a partnership among agencies (Department of Defense, National Oceanic and Atmospheric Administration, NASA, National Science Foundation, Department of Energy) counting on long-term use of ESMF to underpin their scientific modeling research and operational prediction capabilities. These agencies are teaming up to share costs and coordinate long-term planning and development beyond the critical transition point at the end of FY2005, when funding runs out from the NASA Earth-Sun System Technology Office (ESTO) Computational Technologies Project that has supported ESMF development through the present.

Program managers and leads from the various agencies as well as representatives from the ESMF Executive and Advisory Boards participated in the meeting. Don Anderson, program manager for Modeling, Analysis, and Prediction at NASA Headquarters, served as meeting chair. All participating agency leads presented their agencies' strategic objectives involving

Participants in the second meeting of the Earth System Modeling Framework Interagency Working Group, November 30, 2004, at NASA GSFC in Greenbelt, MD (Photo credit: Richard Glassbrook, AMTI).

ESMF. IAWG members heard status updates on ESMF development and applications as well as descriptions of experiments aimed at establishing software interoperability with PRISM, ESMF's European counterpart. The group also discussed formalizing each agency's appointments to the IAWG and officially reviewing a project plan that details a structure for the many stakeholder institutions to collaborate in the evolution of the ESMF software. The draft ESMF Project Plan and technical and programmatic information about ESMF can be found at http://www.esmf.ucar.edu/.

This is the final issue of ESDCD News. The former ESDCD (Code 930) has been reorganized under the Computational & Information Sciences and Technology Office (CISTO, Code 606) and the Software Integration and Visualization Office (SIVO, Code 610.3) of the GSFC Sciences and Exploration Directorate (Code 600).

Look for more information at http://science.gsfc.nasa.gov/